PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : C12N 15/00, 15/09, 15/29, 15/82, A01H 1/00, 3/00, 5/00 — A1

(11) International Publication Number: **WO 98/37184**

(43) International Publication Date: 27 August 1998 (27.08.98)

(21) International Application Number: PCT/US98/02998

(22) International Filing Date: 20 February 1998 (20.02.98)

(30) Priority Data:

08/804,534	21 February 1997 (21.02.97)	US
Not furnished	19 February 1998 (19.02.98)	US

(71) Applicant: THE REGENTS OF THE UNIVERSITY OF CALIFORNIA [-/US]; 22nd floor, 300 Lakeside Drive, Oakland, CA 94612-3550 (US).

(72) Inventors: HARADA, John, J.; 115 Grande Avenue, Davis, CA 95616 (US). LOTAN, Tamar; 408 Russell Park #4, Davis, CA 95616 (US). OHTO, Masa-aki; 2-5-1, Tatsumi-minami #1-23, Okazaki 444 (JP). GOLDBERG, Robert, B.; 21907 Canon Drive, Topanga, CA 90290 (US). FISCHER, Robert, L.; 1423 Scott Street, El Cerrito, CA 94530 (US).

(74) Agent: BERLINER, Robert; Fulbright & Jaworski, L.L.P., 29th floor, 865 S. Figueroa Street, Los Angeles, CA 90017-2571 (US).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, GW, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

Published
With international search report.

(54) Title: LEAFY COTYLEDON1 GENES AND THEIR USES

(57) Abstract

The present invention provides nucleic acid sequences from embryo-specific genes. The nucleic acids are useful in targeting gene expression to embryos or in modulating embryo development.

LEAFY COTYLEDON1 GENES AND THEIR USES

FIELD OF THE INVENTION

The present invention is directed to plant genetic engineering. In particular, it relates to new embryo-specific genes useful in improving agronomically important plants.

BACKGROUND OF THE INVENTION

Embryogenesis in higher plants is a critical stage of the plant life cycle in which the primary organs are established. Embryo development can be separated into two main phases: the early phase in which the primary body organization of the embryo is laid down and the late phase which involves maturation, desiccation and dormancy. In the early phase, the symmetry of the embryo changes from radial to bilateral, giving rise to a hypocotyl with a shoot meristem surrounded by the two cotyledonary primordia at the apical pole and a root meristem at the basal pole. In the late phase, during maturation the embryo achieves its maximum size and the seed accumulates storage proteins and lipids. Maturation is ended by the desiccation stage in which the seed water content decreases rapidly and the embryo passes into metabolic quiescent state. Dormancy ends with seed germination, and development continues from the shoot and the root meristem regions.

The precise regulatory mechanisms which control cell and organ differentiation during the initial phase of embryogenesis are largely unknown. The plant hormone abscisic acid (ABA) is thought to play a role during late embryogenesis, mainly in the maturation stage by inhibiting germination during embryogenesis (Black, M. (1991). In Abscisic Acid: Physiology and Biochemistry, W. J. Davies and H. G. Jones, eds. (Oxford: Bios Scientific Publishers Ltd.), pp. 99-124) Koornneef, M., and Karssen, C. M. (1994). In Arabidopsis, E. M. Meyerowitz and C. R. Sommerville, eds. (Cold Spring Harbor: Cold Spring Harbor Laboratory Press), pp. 313-334). Mutations which effect seed development and are ABA insensitive have been identified in *Arabidopsis* and maize. The ABA insensitive (abi3) mutant of *Arabidopsis* and the viviparous1 (vp1) mutant of maize are detected mainly during late embryogenesis (McCarty, *et al.*, (1989) *Plant Cell* 1, 523-532 and Parcy *et al.*, (1994) *Plant Cell* 6, 1567-1582). Both the *VP1* gene and the *ABI3* genes have been isolated and were

found to share conserved regions (Giraudat, J. (1995) *Current Opinion in Cell Biology* 7:232-238 and McCarty, D. R. (1995). *Annu. Rev. Plant Physiol. Plant Mol. Biol.* 46:71-93). The *VP1* gene has been shown to function as a transcription activator (McCarty, *et al.*, (1991) *Cell* 66:895-906). It has been suggested that *ABI3* has a similar function.

Another class of embryo defective mutants involves three genes: *LEAFY COTYLEDON1* and *2* (*LEC1, LEC2*) and *FUSCA3 (FUS3)*. These genes are thought to play a central role in late embryogenesis (Baumlein, *et al.* (1994) *Plant J.* 6:379-387; Meinke, D. W. (1992) *Science* 258:1647-1650; Meinke *et al., Plant Cell* 6:1049-1064; West *et al.*, (1994) *Plant Cell* 6:1731-1745). Like the abi3 mutant, leafy cotyledon-type mutants are defective in late embryogenesis. In these mutants, seed morphology is altered, the shoot meristem is activated early, storage proteins are lacking and developing cotyledons accumulate anthocyanin. As with abi3 mutants, they are desiccation intolerant and therefore die during late embryogenesis. Nevertheless, the immature mutants embryos can be rescued to give rise to mature and fertile plants. However, unlike abi3 when the immature mutants germinate they exhibit trichomes on the adaxial surface of the cotyledon. Trichomes are normally present only on leaves, stems and sepals, not cotyledons. Therefore, it is thought that the leafy cotyledon type genes have a role in specifying cotyledon identity during embryo development.

Among the above mutants, the *lec1* mutant exhibits the most extreme phenotype during embryogenesis. For example, the maturation and postgermination programs are active simultaneously in the *lec1* mutant (West et al., 1994), suggesting a critical role for *LEC1* in gene regulation during late embryogenesis.

In spite of the recent progress in defining the genetic control of embryo development, further progress is required in the identification and analysis of genes expressed specifically in the embryo and seed. Characterization of such genes would allow for the genetic engineering plants with a variety of desirable traits. For instance, modulation of the expression of genes which control embryo development may be used to alter traits such as accumulation of storage proteins in leaves and cotyledons. Alternatively, promoters from embryo or seed-specific genes can be used to direct expression of desirable heterologous genes to the embryo or seed. The present invention addresses these and other needs.

SUMMARY OF THE INVENTION

The present invention is based, in part, on the isolation and characterization of *LEC1* genes. The invention provides isolated nucleic acid molecules comprising a *LEC1* polynucleotide sequence, typically about 630 nucleotides in length, which specifically hybridizes to SEQ. ID. No. 1 under stringent conditions. The *LEC1* polynucleotides of the invention can encode a LEC1 polypeptide of about 210 amino acids, typically as shown in SEQ. ID. No. 2.

The nucleic acids of the invention may also comprise expression cassettes containing a plant promoter operably linked to the *LEC1* polynucleotide. In some embodiments, the promoter is from a *LEC1* gene, for instance, as shown in SEQ. ID. No. 3. The *LEC1* polynucleotide may be linked to the promoter in a sense or antisense orientation.

The invention also provides transgenic plants comprising an expression cassette containing a plant promoter operably linked to a heterologous *LEC1* polynucleotide. The *LEC1* may encode a LEC1 polypeptide or may be linked to the promoter in an antisense orientation. The plant promoter may be from any number of sources, including a *LEC1* gene, such a as that shown in SEQ. ID. No. 3 or SEQ. ID. No. 4. The transgenic plant can be any desired plant but is often a member of the genus *Brassica*.

Methods of modulating seed development in a plants are also provided. The methods comprise introducing into a plant an expression cassette containing a plant promoter operably linked to a heterologous *LEC1* polynucleotide. The *LEC1* may encode a LEC1 polypeptide or may be linked to the promoter in an antisense orientation. The expression cassette can be introduced into the plant by any number of means known in the art, including through a sexual cross.

The invention further provides expression cassettes containing promoter sequences from *LEC1* genes. The promoters of the invention can be characterized by their ability to specifically hybridizes to a polynucleotide sequence consisting of nucleotides 1 to -1998 of SEQ. ID. No. 3. The promoters of the invention can be operably linked to a variety of nucleic acids, whose expression is to be targeted to embryos or seeds. Transgenic plants comprising the expression cassettes are also provided.

The promoters of the invention can be used in methods of targeting expression of a desired polynucleotide to seeds. The methods comprise introducing into a plant an

expression cassette containing a *LEC1* promoter operably linked to a heterologous polynucleotide sequence.

Definitions

The phrase "nucleic acid" refers to a single or double-stranded polymer of deoxyribonucleotide or ribonucleotide bases read from the 5' to the 3' end. Nucleic acids may also include modified nucleotides that permit correct read through by a polymerase and do not alter expression of a polypeptide encoded by that nucleic acid.

The phrase "polynucleotide sequence" or "nucleic acid sequence" includes both the sense and antisense strands as either individual single strands or in the duplex. It includes, but is not limited to, self-replicating plasmids, chromosomal sequences, and infectious polymers of DNA or RNA.

The phrase "nucleic acid sequence encoding" refers to a nucleic acid which directs the expression of a specific protein or peptide. The nucleic acid sequences include both the DNA strand sequence that is transcribed into RNA and the RNA sequence that is translated into protein. The nucleic acid sequences include both the full length nucleic acid sequences as well as non-full length sequences derived from the full length sequences. It should be further understood that the sequence includes the degenerate codons of the native sequence or sequences which may be introduced to provide codon preference in a specific host cell.

The term "promoter" refers to a region or sequence determinants located upstream or downstream from the start of transcription and which are involved in recognition and binding of RNA polymerase and other proteins to initiate transcription. A "plant promoter" is a promoter capable of initiating transcription in plant cells. Such promoters need not be of plant origin, for example, promoters derived from plant viruses, such as the CaMV35S promoter, can be used in the present invention.

The term "plant" includes whole plants, plant organs (e.g., leaves, stems, flowers, roots, etc.), seeds and plant cells and progeny of same. The class of plants which can be used in the method of the invention is generally as broad as the class of higher plants amenable to transformation techniques, including both monocotyledonous and dicotyledonous plants, as well as certain lower plants such as algae. It includes plants of a variety of ploidy levels, including polyploid, diploid and haploid.

stages of development, while the single mutants developed into mature embryos, suggesting that these genes act early during development.

Further examination of the early stages of the single and double mutations showed defects in the shape, size and cell division pattern of the mutants suspensors. The suspensor of wild type embryo consists of a single file of six to eight cells, whereas the suspensors of the mutants are often enlarged and undergo periclinal divisions. Leafy cotyledon mutants exhibit suspensor anomalies at the globular or transition stage whereas wild type and abi3 mutant do not show any abnormalities.

The number of anomalous suspensors increases as the embryos continue to develop. At the torpedo stage, the wild type suspensor cells undergo programmed cell death, but in the mutants secondary embryos often develop from the abnormal suspensors and, when rescued, give rise to twins.

The Organization of the *LEC1* Locus in Wild Type Plants and *lec1* Mutants

Two mutant alleles of the LEC1 gene have been reported, *lec1-1* and *lec1-2* (Meinke, 1992; West et al., 1994). Both mutants were derived from a population of plants mutagenized insertionally with T-DNA (Feldmann and Marks, 1987), although *lec1-1* is not tagged. The *lec1-2* mutant contains multiple T-DNA insertions. A specific subset of T-DNA fragments were found to be closely linked with the mutation. A genomic library of *lec1-2* was screened using right and left borders T-DNA as probes. Genomic clones containing T-DNA fragments that cosegragate with the mutation were isolated and tested on southern blots of both wild type and *lec1-1* plants. Only one clone hybridized with *Arabidopsis* DNA and also gave polymorphic restriction fragment in *lec1-1*.

The *lec1-1* polymorphism resulted from a small deletion, approximately 2 kb in length. Using sequences from the plant fragment flanking the T-DNA, the genomic wild type DNA clones and the *lec1-1* genomic clones were isolated. An EcoRI fragment of 7.4 kb of the genomic wild type DNA that corresponded to the polymorphic restriction fragment in lec1-1 was further analyzed and sequenced. The exact site of the deletion in *lec1-1* was identified using a PCR fragment that was generated by primers, within the expected borders of the deleted fragment, and sequencing.

In the wild type genomic DNA that corresponded to the *lec1-1* deletion, a 626 bp ORF was identified. Southern analysis of wild type DNA and the two mutants DNA probed with the short DNA fragment of the ORF revealed that both the wild type and *lec1-2*

DNA contain the ORF while the *lec1-1* genomic DNA did not hybridize. The exact insertion site of the T-DNA in *lec1-2* mutant was determined by PCR and sequencing and it was found that the T-DNA was inserted 115 bp upstream of the ORF's translational initiation codon in the 5' region of the gene.

At the site of the T-DNA insertion a small deletion of 21 plant nucleic acids and addition of 20 unknown nucleic acids occurred. These results suggest that in *lec1-2* the T-DNA interferes with the regulation of the ORF while in *lec1-1* the whole gene is deleted. Thus, both *lec1* alleles contain DNA disruptions at the same locus, confirming the identity of the *LEC1* locus.

The *lec1* Mutants Can Be Complement by Transformation

To prove that the 7.4 kb genomic wild type fragment indeed contained the ORF of the *LEC1* gene, we used a genomic fragment of 3395 bp (SEQ. ID. No. 3) within that fragment to transform homozygous *lec1-1* and *lec1-2* plants. The clone consists of a 3395 bp BstYI restriction fragment containing the gene and the promoter region. The translation start codon (ATG) of the polypeptide is at 1999 and the stop codon is at 2625 (TGA). There are no introns in the gene.

The transformed plants were selected on hygromycin plates and were tested to contain the wild type DNA fragment by PCR analysis. Both transgenic mutants were able to produce viable progeny, that were desiccation tolerant and did not posses trichomes on their cotyledons. We concluded that the 3.4 kb fragment can complement the lec1 mutation and since there is only one ORF in the deleted 2 kb fragment in *lec1-1* we suggest that this ORF corresponds to the *LEC1* gene.

The *LEC1* Gene is a Member of Gene Family

In order to isolate the *LEC1* gene two cDNA libraries of young siliques were screened using the 7.4 kb DNA fragment as a probe. Seventeen clones were isolated and after further analysis and partial sequencing they were all found to be identical to the genomic ORF. The cDNA contains 626 bp ORF specifying 208 amino acid protein (SEQ. ID. Nos. 1 and 2).

The *LEC1* cDNA was used to hybridize a DNA gel blot containing Ws-O genomic DNA digested with three different restriction enzymes. Using low stringency hybridization we found that there is at least one more gene. This confirmed our finding of two more Arabidopsis ESTs that show homology to the *LEC1* gene.

The *LEC1* gene is Embryo Specific

The *lec1* mutants are affected mostly during embryogenesis. Rescued mutants can give rise to homozygous plants that have no obvious abnormalities other than the presence of trichomes on their cotyledons and their production of defective progeny. Therefore, we expected the *LEC1* gene to have a role mainly during embryogenesis and not during vegetative growth. To test this assumption Poly (A)+ RNA was isolated from siliques, seedling, roots, leaves, stems and buds of wild type plants and from siliques of lec1 plants. Only one band was detected on northern blots using either the *LEC1* gene as a probe or the 7.4 kb genomic DNA fragment suggesting that there is only one gene in the genomic DNA fragment which is active transcriptionally. The transcript was detected only in siliques containing young and mature embryos and was not detected in seedlings, roots, leaves, stems and buds indicating that the *LEC1* gene is indeed embryo specific. In addition, no RNA was detected in siliques of both alleles of lec1 mutants confirming that this ORF corresponds to the LEC1 gene.

Expression Pattern of the LEC1 Gene

To study how the *LEC1* gene specifies cotyledons identity, we analyzed its expression by in situ hybridization. We specifically focused on young developing embryos since the mutants abnormal suspensors phenotype indicates that the *LEC1* gene should be active very early during development.

During embryogenesis, the *LEC1* transcript was first detected in proglobular embryos. The transcript was found in all cells of the proembryo and was also found in the suspensor and the endosperm. However, from the globular stage and on it accumulates more in the outer layer of the embryo, namely the protoderm and in the outer part of the ground meristem leaving the procambium without a signal. At the torpedo stage the signal was stronger in the cotyledons and the root meristem, and was more limited to the protoderm layer. At the bent cotyledon stage the signal was present throughout the embryo and at the last stage of development when the embryo is mature and filling the whole seed we could not detect the *LEC1* transcript. This might be due to sensitivity limitation and may imply that if the *LEC1* transcript is expressed at that stage it is not localized in the mature embryo, but rather spread throughout the embryo.

The *LEC1* gene encodes a Homolog of CCAAT binding factor.

WHAT IS CLAIMED IS:

1. An isolated nucleic acid molecule comprising a *LEC1* polynucleotide sequence, which polynucleotide sequence specifically hybridizes to SEQ. ID. No. 1 under stringent conditions.

2. The isolated nucleic acid molecule of claim 1, wherein the *LEC1* polynucleotide is between about 100 nucleotides and about 630 nucleotides in length.

3. The isolated nucleic acid molecule of claim 1, wherein the *LEC1* polynucleotide is SEQ. ID. No. 1.

4. The isolated nucleic acid molecule of claim 1, wherein the *LEC1* polynucleotide encodes a LEC1 polypeptide of between about 50 and about 210 amino acids.

5. The isolated nucleic acid molecule of claim 4, wherein the LEC1 polypeptide has an amino acid sequence as shown in SEQ. ID. No. 2.

6. The isolated nucleic acid molecule of claim 1, further comprising a plant promoter operably linked to the *LEC1* polynucleotide.

7. The isolated nucleic acid molecule of claim 6, wherein the plant promoter is from a *LEC1* gene.

8. The isolated nucleic acid of claim 7, wherein the *LEC1* gene is as shown in SEQ. ID. No. 3.

9. The isolated nucleic acid of claim 7, wherein the *LEC1* gene is as shown in SEQ. ID. No. 4.

10. The isolated nucleic acid of claim 7, wherein the *LEC1* polynucleotide is linked to the promoter in an antisense orientation.

11. An isolated nucleic acid molecule comprising a *LEC1* polynucleotide sequence, which polynucleotide sequence encodes LEC1 polypeptide of between about 50 and about 210 amino acids.

12. The isolated nucleic acid of claim 10, wherein the LEC1 polypeptide has an amino acid sequence as shown in SEQ. ID. No. 2.

13. A transgenic plant comprising an expression cassette containing a plant promoter operably linked to a heterologous *LEC1* polynucleotide that specifically hybridizes to SEQ. ID. No. 1 under stringent conditions.

14. The transgenic plant of claim 12, wherein the heterologous *LEC1* polynucleotide encodes a LEC1 polypeptide.

15. The transgenic plant of claim 13, wherein the LEC1 polypeptide is SEQ. ID. No. 2.

16. The transgenic plant of claim 12, wherein the heterologous *LEC1* polynucleotide is linked to the promoter in an antisense orientation.

17. The transgenic plant of claim 12, wherein the plant promoter is from a *LEC1* gene.

18. The transgenic plant of claim 16, wherein the *LEC1* gene is as shown in SEQ. ID. No. 3.

19. The transgenic plant of claim 12, which is a member of the genus *Brassica*.

20. A method of modulating seed development in a plant, the method comprising introducing into the plant an expression cassette containing a plant promoter

operably linked to a heterologous *LEC1* polynucleotide that specifically hybridizes to SEQ. ID. No. 1 under stringent conditions.

21. The method of claim 19, wherein the heterologous *LEC1* polynucleotide encodes a LEC1 polypeptide.

22. The method of claim 20, wherein the LEC1 polypeptide has an amino acid sequence as shown in SEQ. ID. No. 2.

23. The method of claim 19, wherein the heterologous *LEC1* polynucleotide is linked to the promoter in an antisense orientation.

24. The method of claim 19, wherein the heterologous *LEC1* polynucleotide is SEQ. ID. No. 1.

25. The method of claim 19, wherein the plant promoter is from a *LEC1* gene.

26. The method of claim 19, wherein the *LEC1* gene is as shown in SEQ. ID. No. 3.

27. The method of claim 19, wherein the plant is a member of the genus *Brassica*.

28. The method of claim 19, wherein the expression cassette is introduced into the plant through a sexual cross.

29. An isolated nucleic acid molecule comprising a plant promoter that specifically hybridizes to a polynucleotide sequence consisting of nucleotides 1 to 1998 of SEQ. ID. No. 3.

30. The isolated nucleic acid molecule of claim 28, wherein the plant promoter sequence consists essentially of nucleotides 1 to 1998 of SEQ. ID. No. 3.

31. The isolated nucleic acid molecule of claim 28, wherein the plant promoter sequence is a subsequence of SEQ. ID. No. 4.

32. The isolated nucleic acid molecule of claim 28, further comprising a polynucleotide sequence operably linked to the plant promoter sequence.

33. The isolated nucleic acid of claim 30, wherein the polynucleotide sequence operably linked to the plant promoter sequence encodes a desired polypeptide.

34. The isolated nucleic acid molecule of claim 28, wherein the polynucleotide sequence is linked to the promoter in an antisense orientation.

35. A transgenic plant comprising an expression cassette containing a *LEC1* promoter operably linked to a heterologous polynucleotide sequence, wherein the *LEC1* promoter specifically hybridizes to SEQ. ID. No. 3 under stringent conditions.

36. The transgenic plant of claim 33, wherein the polynucleotide sequence encodes a desired polypeptide.

37. The transgenic plant of claim 33, wherein the heterologous polynucleotide sequence is linked to the *LEC1* promoter in an antisense orientation.

38. The transgenic plant of claim 33, wherein the *LEC1* promoter is as shown in SEQ. ID. No. 3.

39. The transgenic plant of claim 33, which is a member of the genus *Brassica*.

40. A method of targeting expression of a polynucleotide to a seed, the method comprising introducing into a plant an expression cassette containing a *LEC1* promoter operably linked to a heterologous polynucleotide sequence, wherein the *LEC1* promoter specifically hybridizes to a polynucleotide sequence consisting of nucleotides 1 to -1998 of SEQ. ID. No. 3.

41. The method of claim 38, wherein the heterologous polynucleotide sequence encodes a desired polypeptide.

42. The method of claim 38, wherein the heterologous polynucleotide sequence is linked to the promoter in an antisense orientation.